BU #12



PCT09

# **ENTERED**

RAW SEQUENCE LISTING

4 <110> APPLICANT: CHEN, WENFANG

PATENT APPLICATION: US/09/889,344B

DATE: 12/04/2002 TIME: 14:00:04

00.F1 .dn11

Input Set : A:\USSEQLIST3.txt

Output Set: N:\CRF4\12032002\I889344B.raw

```
MEEK, THOMAS D.
      5
      6
              POWELL, DAVID J.
      7
              TEW, DAVID G.
     10 <120> TITLE OF INVENTION: Method of Site Specific Labeling of Proteins and Uses
              Therefor
     14 <130> FILE REFERENCE: P50892
     16 <140> CURRENT APPLICATION NUMBER: 09/889,344B
     17 <141> CURRENT FILING DATE: 2001-07-16
     19 <150> PRIOR APPLICATION NUMBER: PCT/US00/01481
     20 <151> PRIOR FILING DATE: 2000-01-20
     22 <150> PRIOR APPLICATION NUMBER: US 60/117,327
     23 <151> PRIOR FILING DATE: 1999-01-22
     25 <160> NUMBER OF SEQ ID NOS: 16
     27 <170> SOFTWARE: FastSEQ for Windows Version 3.0
     29 <210> SEQ ID NO: 1
     30 <211> LENGTH: 5
     31 <212> TYPE: PRT
     32 <213> ORGANISM: Artificial Sequence
     34 <220> FEATURE:
     35 <221> NAME/KEY: unsure
     36 <222> LOCATION: (5)
     37 <223> OTHER INFORMATION: Where Xaa at position (5) can represent Leucine or Isoleucine
     39 <220> FEATURE:
     40 <223> OTHER INFORMATION: Wherein the amino acid sequence is modified by reacting a
     41 <223> OTHER INFORMATION: transglutaminase with a detectable labeling compound
W--> 42 <400> 1
W--> 43 Gln Ser Lys Val Xaa
     44
          1
     46 <210> SEQ ID NO: 2
     47 <211> LENGTH: 207
     48 <212> TYPE: PRT
     49 <213> ORGANISM: Artificial Sequence
     51 <220> FEATURE:
     52 <221> NAME/KEY: unsure
     53 <222> LOCATION: (1)(2)(3)(4)(5)(6)(7)(8)(9)(10)(11)(12)(13)(14)(15)
     54 <222> LOCATION: (16)(17)(18)(19)(20)(21)(22)(23)(24)(25)(26)(27)(28)
W--> 55 <222> (29) (30) (31) (32) (33) (34) (35) (36) (37) (38) (39) (40) (41)
W--> 56 <222> (42) (43) (44) (45) (46) (47) (48) (49) (50) (51) (52) (53) (54)
W--> 57 <222> (55) (56) (57) (58) (59) (60) (61) (62) (62) (64) (65) (66) (67)
W-->58 <222> (68) (69) (70) (71) (72) (73) (74) (75) (76) (77) (78) (79) (80)
W--> 59 <222> (81) (82) (83) (84) (85) (86) (87) (88) (89) (90) (91) (92) (93)
W--> 60 <222> (94) (95) (96) (97) (98) (99) (100) (101) (106) (107) (108) (109)
```

#### RAW SEQUENCE LISTING DATE: 12/04/2002 PATENT APPLICATION: US/09/889,344B TIME: 14:00:05 Input Set : A:\USSEQLIST3.txt Output Set: N:\CRF4\12032002\I889344B.raw W--> 61 <222> (110) (111) (112) (113) (114) (115) (116) (117) (118) (119) (120) W--> 62 <222> (121) (122) (123) (124) (125) (126) (127) (128) (129) (130) (131) W--> 63 <222> (132) (133) (134) (135) (136) (137) (138) (139) (140) (141) (142) W--> 64 <222> (143) (144) (145) (146) (147) (148) (149) (150) (151) (152) (153) W--> 65 <222> (154) (155) (156) (157) (158) (159) (160) (161) (162) (163) (164) W--> 66 <222> (165) (166) (167) (168) (169) (170) (171) (172) (173) (174) (175) W--> 67 <222> (176) (177) (178) (179) (180) (181) (182) (183) (184) (185) (186) W--> 68 <222> (187) (188) (189) (190) (191) (192) (193) (194) (195) (196) (197) W--> 69 <222> (198) (199) (200) (201) (202) (203) (204) (205) (206) (207) 70 <223> OTHER INFORMATION: Where Xaa can represent none or any one of the twenty naturally 71 <223> OTHER INFORMATION: occurring amino acids W--> 73 <220> FEATURE: 74 <223> OTHER INFORMATION: Wherein the amino acid sequence is modified by reacting a 75 <223> OTHER INFORMATION: transglutaminase with a detectable labeling compound W--> 77 < 400 > 279 5 10 W--> 80 81 25 W--> 82 83 40 W--> 84 85 W--> 86 87 75 70 W--> 88 89 90 W--> 90 Xaa Xaa Xaa Xaa Gln Ser Lys Val Xaa Xaa Xaa Xaa Xaa Xaa Xaa 91 100 105 W--> 92 93 115 120 W--> 94 95 130 135 140 W--> 96 97 W--> 9899 165 170 W--> 100 101 180 185 W--> 102 103 195 200 105 <210> SEQ ID NO: 3 106 <211> LENGTH: 207

112 <222> LOCATION: (1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11) (12) (13) (14) (15) (13 <222> LOCATION: (16) (17) (18) (19) (20) (21) (22) (23) (24) (25) (26) (27) (28)

111 <221> NAME/KEY: unsure

107 <212> TYPE: PRT

110 <220> FEATURE:

108 <213> ORGANISM: Artificial Sequence



### RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/889,344B

DATE: 12/04/2002 TIME: 14:00:05

Input Set : A:\USSEQLIST3.txt

Output Set: N:\CRF4\12032002\I889344B.raw

- W--> 114 <222> (29) (30) (31) (32) (33) (34) (35) (36) (37) (38) (39) (40) (41)
- W--> 115 < 222> (42) (43) (44) (45) (46) (47) (48) (49) (50) (51) (52) (53) (54)
- W--> 116 < 222> (55) (56) (57) (58) (59) (60) (61) (62) (62) (64) (65) (66) (67)
- W--> 117 <222> (68) (69) (70) (71) (72) (73) (74) (75) (76) (77) (78) (79) (80)
- W--> 118 < 222> (81) (82) (83) (84) (85) (86) (87) (88) (89) (90) (91) (92) (93)
- W--> 119 <222> (94) (95) (96) (97) (98) (99) (100) (101) (106) (107) (108) (109)
- W--> 120 <222> (110) (111) (112) (113) (114) (115) (116) (117) (118) (119) (120)
- W--> 121 <222> (121) (122) (123) (124) (125) (126) (127) (128) (129) (130) (131)
- W--> 122 <222> (132) (133) (134) (135) (136) (137) (138) (139) (140) (141) (142)
- W--> 123 <222> (143) (144) (145) (146) (147) (148) (149) (150) (151) (152) (153)
- W--> 124 < 222> (154) (155) (156) (157) (158) (159) (160) (161) (162) (163) (164)
- W--> 125 < 222> (165) (166) (167) (168) (169) (170) (171) (172) (173) (174) (175)<math>W--> 126 < 222> (176) (177) (178) (179) (180) (181) (182) (183) (184) (185) (186)
- W--> 126 < 222> (176) (177) (178) (179) (180) (181) (182) (183) (184) (185) (186) (187) (188) (189) (190) (191) (192) (193) (194) (195) (196) (197)
- W--> 128 < 222> (198) (199) (200) (201) (202) (203) (204) (205) (206) (207)
- 129 < 223 > OTHER INFORMATION: Where Xaa can represent none or any one of the twenty naturally
- 130 <223> OTHER INFORMATION: occurring amino acids
- W--> 132 <220> FEATURE:
  - 133 <223> OTHER INFORMATION: Wherein the amino acid sequence is modified by reacting a
    - 134 <223> OTHER INFORMATION: transglutaminase with a detectable labeling compound
- W--> 136 <400> 3

W--> 139

- - 140 20 25 30
- - 142 35 40 4
- - 146 65 70 75 80
- W--> 149 Xaa Xaa Xaa Xaa Xaa Gln Ser Lys Val Xaa Xaa Xaa Xaa Xaa Xaa Xaa 150 100 105 110
- 152 115 120 125

- - 164 <210> SEQ ID NO: 4
  - 165 <211> LENGTH: 10

## RAW SEQUENCE LISTING DATE: 12/04/2002 PATENT APPLICATION: US/09/889,344B TIME: 14:00:05

Input Set : A:\USSEQLIST3.txt

Output Set: N:\CRF4\12032002\I889344B.raw

```
166 <212> TYPE: PRT
167 <213> ORGANISM: Artificial Sequence
169 <220> FEATURE:
170 <223> OTHER INFORMATION: Derivative of a factor XIII substrate
172 <400> SEQUENCE: 4
173 Leu Ser Leu Ser Gln Ser Lys Val Leu Gly
176 <210> SEQ ID NO: 5
177 <211> LENGTH: 10
178 <212> TYPE: PRT
179 <213> ORGANISM: Artificial Sequence
181 <220> FEATURE:
182 <223> OTHER INFORMATION: Derivative of a factor XIII substrate
184 <400> SEQUENCE: 5
185 Ile Gly Glu Gly Gln Ser Lys Val Leu Gly
186
188 <210> SEO ID NO: 6
189 <211> LENGTH: 10
190 <212> TYPE: PRT
191 <213> ORGANISM: Artificial Sequence
193 <220> FEATURE:
194 <223> OTHER INFORMATION: Derivative of a factor XIII substrate
196 <400> SEQUENCE: 6
197 Leu Gly Pro Gly Gln Ser Lys Val Ile Gly
198
    1
200 <210> SEQ ID NO: 7
201 <211> LENGTH: 81
202 <212> TYPE: DNA
203 <213> ORGANISM: Unknown
205 <220> FEATURE:
206 <223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag
208 <400> SEQUENCE: 7
209 tgtacctcag accatatgag cctgtccctg tcccagtcca aagttctgcc gggtccgagc
                                                                             60
210 actatcgaag aacgcgttaa g
                                                                             81
212 <210> SEQ ID NO: 8
213 <211> LENGTH: 37
214 <212> TYPE: DNA
215 <213> ORGANISM: Unknown
217 <220> FEATURE:
218 <223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag
220 <400> SEQUENCE: 8
221 tgatgtcagt caagcttacg cctggtggcc gttgatg
                                                                             37
223 <210> SEQ ID NO: 9
224 <211> LENGTH: 14
225 <212> TYPE: PRT
226 <213> ORGANISM: Artificial Sequence
228 <220> FEATURE:
229 <223> OTHER INFORMATION: Derivative of a factor XIII substrate
231 <400> SEQUENCE: 9
```

### RAW SEQUENCE LISTING DATE: 12/04/2002 PATENT APPLICATION: US/09/889,344B TIME: 14:00:05

Input Set : A:\USSEQLIST3.txt

Output Set: N:\CRF4\12032002\1889344B.raw

232 Met Ser Leu Ser Leu Ser Gln Ser Lys Val Leu Pro Gly Pro 233 10 235 <210> SEO ID NO: 10 236 <211> LENGTH: 37 237 <212> TYPE: DNA 238 <213> ORGANISM: Unknown 240 <220> FEATURE: 241 <223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag 243 <400> SEQUENCE: 10 37 244 tgtacctcag accatatgag cactatcgaa gaacgcg 246 <210> SEQ ID NO: 11 247 <211> LENGTH: 78 248 <212> TYPE: DNA 249 <213> ORGANISM: Unknown 251 <220> FEATURE: 252 <223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag 254 <400> SEQUENCE: 11 255 tgatgtcagt caagettacg gacceggeag aactttggac tgggacaggg acagegeetg 60 256 gtggccgttg atgtaatc 78 258 <210> SEQ ID NO: 12 259 <211> LENGTH: 12 260 <212> TYPE: PRT 261 <213> ORGANISM: Artificial Sequence 263 <220> FEATURE: 264 <223> OTHER INFORMATION: Derivative of E. coli ACP protein 266 <400> SEQUENCE: 12 267 Leu Ser Leu Ser Gln Ser Lys Val Leu Pro Gly Pro 268 270 <210> SEQ ID NO: 13 271 <211> LENGTH: 92 272 <212> TYPE: DNA 273 <213> ORGANISM: Unknown 275 <220> FEATURE: 276 <223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag into Streptococcus haemophilus FabH gene 279 <400> SEQUENCE: 13 280 tatcatatga gcctgtccct gtcccagtcc aaagttctgc cgggtccggg taccctcgag 60 281 ggatccgctt ttgcaaaaat aagtcaggtt gc 92 283 <210> SEQ ID NO: 14 284 <211> LENGTH: 53 285 <212> TYPE: DNA 286 <213> ORGANISM: Unknown 288 <220> FEATURE: 289 <223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag into 290 Streptococcus haemophilus FabH gene 292 <400> SEQUENCE: 14 293 ctcagatctg agctcactag tggatcctta aattgtaaga atgagcgtgc ccc 53 295 <210> SEQ ID NO: 15 296 <211> LENGTH: 364

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 12/04/2002 PATENT APPLICATION: US/09/889,344B TIME: 14:00:06

Input Set : A:\USSEQLIST3.txt

Output Set: N:\CRF4\12032002\I889344B.raw

### Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

```
Seq#:1; Xaa Pos. 5
Seq#:2; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22
Seq#:2; Xaa Pos. 23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41
Seq#:2; Xaa Pos. 42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60
Seq#:2; Xaa Pos. 61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79
Seq#:2; Xaa Pos. 80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98
Seg#:2; Xaa Pos. 99,100,101,106,107,108,109,110,111,112,113,114,115,116,117
Seq#:2; Xaa Pos. 118,119,120,121,122,123,124,125,126,127,128,129,130,131
Seq#:2; Xaa Pos. 132,133,134,135,136,137,138,139,140,141,142,143,144,145
Seq#:2; Xaa Pos. 146,147,148,149,150,151,152,153,154,155,156,157,158,159
Seq#:2; Xaa Pos. 160,161,162,163,164,165,166,167,168,169,170,171,172,173
Seq#:2; Xaa Pos. 174,175,176,177,178,179,180,181,182,183,184,185,186,187
Seq#:2; Xaa Pos. 188,189,190,191,192,193,194,195,196,197,198,199,200,201
Seq#:2; Xaa Pos. 202,203,204,205,206,207
Seq#:3; Xaa Pos. 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22
Seq#:3; Xaa Pos. 23,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40,41
Seq#:3; Xaa Pos. 42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58,59,60
Seq#:3; Xaa Pos. 61,62,63,64,65,66,67,68,69,70,71,72,73,74,75,76,77,78,79
Seq#:3; Xaa Pos. 80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95,96,97,98
Seq#:3; Xaa Pos. 99,100,101,106,107,108,109,110,111,112,113,114,115,116,117
Seq#:3; Xaa Pos. 118,119,120,121,122,123,124,125,126,127,128,129,130,131
Seq#:3; Xaa Pos. 132,133,134,135,136,137,138,139,140,141,142,143,144,145
Seq#:3; Xaa Pos. 146,147,148,149,150,151,152,153,154,155,156,157,158,159
Seq#:3; Xaa Pos. 160,161,162,163,164,165,166,167,168,169,170,171,172,173
Seq#:3; Xaa Pos. 174,175,176,177,178,179,180,181,182,183,184,185,186,187
Seq#:3; Xaa Pos. 188,189,190,191,192,193,194,195,196,197,198,199,200,201
Seg#:3; Xaa Pos. 202,203,204,205,206,207
```

### VERIFICATION SUMMARY

DATE: 12/04/2002 PATENT APPLICATION: US/09/889,344B TIME: 14:00:06

Input Set : A:\USSEQLIST3.txt

Output Set: N:\CRF4\12032002\I889344B.raw

```
L:42 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:1
L:42 M:283 W: Missing Blank Line separator, <400> field identifier
L\!:\!43 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:0
L:55 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:56 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:57 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:58 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:59 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:60 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:61 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:62 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:63 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:64 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:65 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:66 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:67 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2 L:68 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:69 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:77 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:2
L:78 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:0
L:80 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:16
L:82 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:32
L:84 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:48
L:86 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:64
L:88 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:80
L:90 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:96
L:92 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:112
L:94 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:128
L:96 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:144
L:98 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:160
L:100 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:176
L:102 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:192
L:114 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:115 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:116 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:117 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:118 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:119 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:120 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3 L:121 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:122 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:123 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:124 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:125 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:126 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:127 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3 L:128 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
L:136 M:258 W: Mandatory Feature missing, <220> Tag not found for SEQ ID#:3
```

DATE: 12/04/2002

#### VERIFICATION SUMMARY

PATENT APPLICATION: US/09/889,344B TIME: 14:00:06

Input Set : A:\USSEQLIST3.txt

Output Set: N:\CRF4\12032002\1889344B.raw

```
L:137 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0
L:139 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:16
L:141 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:32
L:143 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:48
L:145 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:64
L:147 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:80
L:149 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:96
L:151 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:112
L:153 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:128
L:155 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:128
L:157 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:144
L:157 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:160
L:159 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:176
L:161 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:176
```